MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2015/2016

TPR2251 – PATTERN RECOGNITION

(All sections / Groups)

6 October 2015
2.30 pm – 4.30 pm
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of 5 pages with 6 questions only.

2. Attempt **FIVE** out of **SIX** questions. All questions carry equal marks and the distribution of the marks for each question is given.

3. Please print all your answers in the answer booklet provided.

QUESTION 1

(a) Consider a pattern recognition problem to separate fruits in a fruit packaging company. The company packages two types of fruits, namely durian and cempedak (jackfruit). The fruits come to the boxing area on a conveyer belt, and they have to be sorted into their corresponding boxes.

(i) Describe the sensor that can be used for this problem. [1 mark]
(ii) Explain the pre-processing and segmentation processes involved. [2 marks]
(iii) Name FOUR features that may separate the fruits. [1 mark]

(b) Suppose you have been assigned a medical image classification task. You need to classify a given medical image as benign (class A) or malignant (class B). You have access to an image database with a number of patterns, some of which are known to originate from class A and some from class B. You identify two measurable quantities (features), x_1 and x_2, from these images. Figure 1.1 shows a plot of x_1 versus x_2. Each point corresponds to a different image from the database where the circles and triangles denote class A and class B, respectively. Do you think x_1 and x_2 are good feature representations for the classification task? Justify your answer. Use a diagram if necessary. [2 marks]

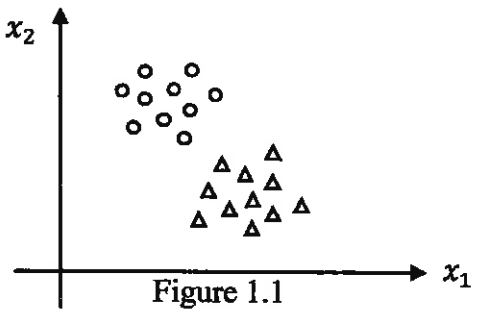


Figure 1.1

(c) Recently there is an outbreak of Middle East Respiratory Syndrome (MERS). A special treatment unit has been setup in the Kuala Lumpur General Hospital to provide special care for the infected patients. Seventy percent of the patients admitted to the special treatment unit have recovered within one month. Of the patients who have recovered, 60% have gastrointestinal symptom; whereas 90% of the patients who have not recovered do not have such symptom. Suppose a patient is diagnosed with MERS, what is the probability that he will recover within one month? [4 marks]

Continued .....

QUESTION 2

(a) Describe a prior probability in Bayes Decision theory. Use the fruit recognition example in Question 1a to illustrate how a prior probability can be formulated for this problem.

[3 marks]

(b) Consider the same problem in Question 1a and answer the following.

(i) Assume the length of the fruit, x, is selected as the feature. The classes for durian and cempedak are denoted as ω_1 and ω_2, respectively. Derive a decision rule for the fruit recognition problem based on the Likelihood Ratio Test assuming equal priors and the class conditional densities,

$$P(x|\omega_1) = \frac{1}{\sqrt{2\pi}} e^{\frac{(x-3)^2}{2}}, P(x|\omega_2) = \frac{1}{\sqrt{2\pi}} e^{\frac{(x-9)^2}{2}}$$

[4 marks]

(ii) How would the decision rule be affected if the prior for durian increases (suppose it is a durian season)?

[1 mark]

(c) True or False: If we train a Naive Bayes classifier using infinite training data that satisfies all of its modeling assumptions (e.g., conditional independence), then it will achieve zero training error over these training examples. Explain your answer.

[2 marks]

Continued .....

QUESTION 3

(a) Answer the following questions based on Bayesian Decision Theory.

(i) Describe a loss function. [1 mark]

(ii) Consider a classification problem working with a d-component feature vector $\mathbf{x}$. Assume there are c classes $\omega_1, \omega_2, \ldots, \omega_c$ and let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be a set of k possible actions. Given that $\lambda(\alpha_i|\omega_j)$ describes the loss incurred for taking action α_i while the state of nature is ω_j, give the expected loss associated with taking action α_i in terms of class-conditional probability densities of $\mathbf{x}$. [1 mark]

(iii) Suppose a bank classifies customers as either good or bad credit risks, denoted as ω_1 and ω_2, respectively. On the basis of extensive historical data, the bank has observed that 60% of the customers are of good credit risks while 40% of bad credit risks. On the basis of a check with a credit bureau, the bank would decide whether to approve a loan request from a customer. The cost (loss) matrix, $\lambda_{i,j}$, for the bank's decision are given as follows:

	ω_1	ω_2
α_1	5	20
α_2	10	0

where α_1 and α_2 denote the decision of the bank to approve and disapprove the loan request, respectively. Calculate the expected loss of choosing α_1 and α_2.

[4 marks]

(b) Consider a two-dimensional Gaussian probability density function, $N(m, S)$, with mean $m = [0,0]^T$ and covariance matrix $S = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$. Draw/sketch the graphs for the Gaussian density function for the following cases. Explain your answers.

(i) $\sigma_1^2 = \sigma_2^2 = 1,\ \sigma_{12} = 0$

(ii) $\sigma_1^2 = 0.2,\ \sigma_2^2 = 2,\ \sigma_{12} = 0$

[4 marks]

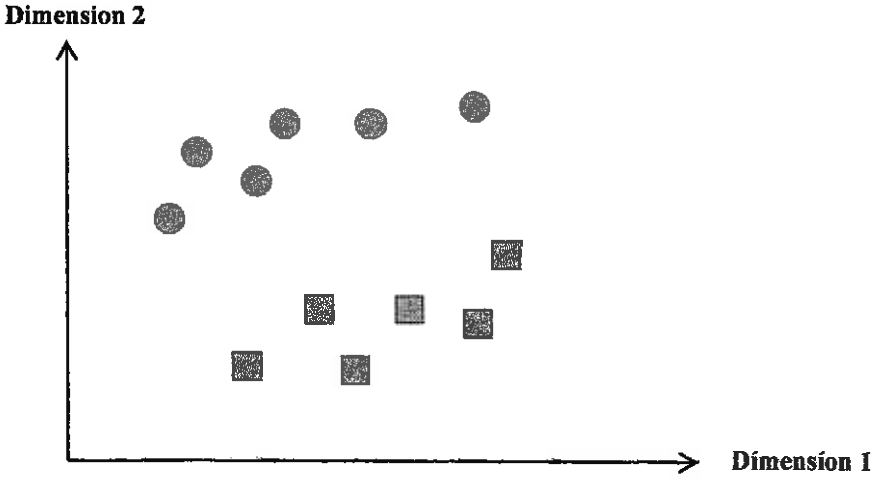
Continued .....

QUESTION 4

(a) Given a data set $X \in \mathbb{R}^N$, describe how principal component analysis (PCA) can be used to reduce the dimension of X to m, where $m < N$.

[2 marks]

(b) Consider the following scattergrams with the *circles and rectangles denoting* class 1 and class 2, respectively. Draw the 1-D subspace (a line) when you project the data onto the subspace of (i) Principal Component Analysis, and (ii) Fisher Linear Discriminant. Provide justification to your answer.



[4 marks]

(c) Given a data set $X = \{2, 2.5, 3, 1, 6\}$, use Parzen windows to estimate the density $p(x)$ at $y = 3$ and $y = 2$. Use the Gaussian function with $\sigma = 1$ as the window function.

[4 marks]

Continued .....

QUESTION 5

(a) Let $x_1, x_2, \ldots, x_N$ be vectors stemmed from a normal distribution with known covariance matrix and unknown mean, that is,

$$p(x_k; \mu) = \frac{1}{(2\pi)^{l/2}|\sum|^{1/2}} \exp\left(-\frac{1}{2}(x_k - \mu)^T \sum^{-1}(x_k - \mu)\right)$$

Obtain the maximum-likelihood estimate of the unknown mean vector. [2 marks]

(b) Describe clustering in pattern recognition and give TWO applications of cluster analysis. [3 marks]

(c) Consider an application of k-means clustering algorithm to classify a two-dimensional dataset $x_1 = (3,5)$, $x_2 = (6,8)$, $x_3 = (4,7)$, $x_4 = (2,5)$, $x_5 = (6,9)$, $x_6 = (8,11)$, $x_7 = (7,8)$, and $x_8 = (6,3)$ into three clusters.

(i) Suppose the initial seeds are x_1, x_2, and x_6. Run the k-means algorithm for 1 epoch and find the new clusters at the end of this epoch. [4 marks]

(ii) Calculate the centers of the new clusters. [1 mark]

QUESTION 6

(a) Explain the starting and stopping criteria for the agglomerative and divisive clustering techniques. [3 marks]

(b) Artificial neural networks (ANNs) have some interesting properties that make these family of machine learning algorithms very appealing when confronting difficult pattern-discovery tasks. In your opinion, what makes ANNs useful when dealing with these difficult task? [3 marks]

(c) Consider a problem to classify four points in a two-dimensional space. Points (-1, 0) and (0, 1) belong to class ω_1; while points (0, -1), and (1, 0) belong to class ω_2. You use a single-layer feed-forward network for this problem. The parameter $\propto$ is set equal to one, and the initial weight vector is chosen as $w(0) = [0,0,0]^T$. Show, step by step, how the error correction learning algorithm adjusts the weights in the perceptron. [4 marks]

End of Page